

On an epidemic model on finite graphs ^{*}

I. Benjamini ¹, L.R. Fontes ², J. Hermon ³, F.P. Machado ²

October 17, 2016

¹Department of Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel.
e-mail: itai.benjamini@weizmann.ac.il.

²Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010,
CEP 05508-900, São Paulo SP, Brasil.
e-mail: lrenato@ime.usp.br, fmachado@ime.usp.br

³Department of Statistics, UC Berkeley. e-mail: jonathan.hermon@berkeley.edu.

Abstract

We study a system of random walks, known as the frog model, starting from an independent $\text{Poisson}(\lambda)$ particle's profile with one additional active particle planted at some vertex \mathbf{o} of a finite connected simple graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Initially, only the particles occupying \mathbf{o} are active. Active particles perform $t \in \mathbb{N} \cup \{\infty\}$ steps of the walk they picked before vanishing and activate all inactive particles they hit. This system is often taken as a model for the spread of an epidemic over a population. Let \mathcal{R}_t be the set of vertices which are visited by the process, when active particles vanish after t steps. We study the susceptibility of the process on the underlying graph, defined as the random quantity $\mathcal{S}(\mathcal{G}) := \inf\{t : \mathcal{R}_t = \mathcal{V}\}$, the time it takes for the entire population get infected. We consider the cases that the underlying graph is either a regular expander or a d -dimensional torus of side length n (for all $d \geq 1$) and determine asymptotic bounds for \mathcal{S} up to a constant factor.

Keywords: frog model, simple random walks, susceptibility, cover times.

^{*}The authors are thankful to Microsoft and CNPq for financial support.

1 Introduction

We study a system of branching random walks known as the *frog model*. The model is often interpreted as a model for a spread of an epidemic or a rumor. The frog model on infinite graphs received much attention, e.g. [19, 1, 2, 18, 9, 10, 12]. As we soon explain in more details, we study a natural parameter associated with the frog model on finite graphs, which in the aforementioned interpretation of the model is meant to capture “how long should a virus live in order to infect the entire population” or “how interesting should a rumor be, so that eventually everybody will hear it”.

Most of the existing literature on the model is focused on the case that the underlying graph on which the particles perform their random walks is \mathbb{Z}^d (for some $d \geq 1$). Beyond the Euclidean setup, there has been much interest in understanding the behavior of the model in the case that the underlying graph is a d -ary tree, either finite or infinite [9, 10, 12]. However, to the best of the authors’ knowledge, the only existing paper concerning the model on finite graphs is [12]. This paper is closely related to [12] (see § 3.3) and also to a paper by the first and third authors about the intimately related random walks social network model (see § 3.4). In this paper we study the model in the case that the underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is some finite connected simple undirected graph. More specifically, we focus mainly on the cases that \mathcal{G} is a d -dimensional torus ($d \geq 1$) of side length n or a regular expander.

The frog model on \mathcal{G} with density λ can be described as follows. Initially there are $\text{Pois}(\lambda)$ particles at each vertex of \mathcal{G} , independently. A site of \mathcal{G} is singled out and called its *origin*, denoted by \mathbf{o} . An additional particle, denoted by w_{plant} , is planted at \mathbf{o} (this is done in order to ensure that the process does not instantly die out). Initially, each particle independently “picks” an infinite trajectory, which is distributed as a discrete-time simple random walk (SRW) on \mathcal{G} started at the particle’s initial position. All particles are inactive (sleeping) at time zero, except for those occupying the origin. Each active particle performs the first τ steps of the walk it picked (for some $\tau \in \mathbb{N} \cup \{\infty\}$) on the vertices of \mathcal{G} (i.e. for τ steps, at each step it moves to a random neighbor of its current position, chosen from the uniform distribution over the neighbor set) after which it cannot become reactivated (one may consider that they vanish). We refer to τ as the particles’ *life time*. Up to the time a particle dies (i.e. during the τ steps of its walk), it activates all sleeping particles it hits along its way. From the moment an

inactive particle is activated, it performs the same dynamics over its life time τ , independently of everything else. We denote the corresponding probability measure by \mathbb{P}_λ . Notice that there is no interaction between active particles, which means that, once activated, each active particle moves independently of everything else.

We now define two natural parameters for the frog model on a finite graph \mathcal{G} . Note that (in contrast to the setup in which \mathcal{G} is infinite) a.s. there exists a finite minimal lifetime τ (which is a function of the initial configuration of the particles and the walks they pick) for which every vertex is visited by an active particle before the process “dies out”. We define this lifetime as $\mathcal{S}(\mathcal{G})$, the *susceptibility* of \mathcal{G} . A more explicit definition of the susceptibility is given in (3.3).

The name Frog model was coined in 1996 by Rick Durrett. It is a particular case of the $A + B \rightarrow 2B$ family of models (see § 3.4). Like other models in this family (e.g. [13, 14]), it is often motivated as a model for the spread of a rumor or infection (e.g. [2]). Keeping this interpretation in mind, the susceptibility is indeed a natural quantity. It is the minimal lifetime τ of a virus (more precisely, of an individual infected by a virus), sufficient for “wiping out the entire population”. In this interpretation, the smaller the law of $\mathcal{S}(\mathcal{G})$ is (stochastically), the more susceptible the population is.

The paper is organized as follows. In Section 2 we present our main results and some conjectures that we believe may drive future research in this subject. In Section 3 we introduce some notations necessary for a formal construction of the model, present a concise introduction to the topic Frog model on trees [12], introduce some related models and examples besides auxiliary results handy for the proofs of the main results of the paper, which are delivered in Section 4.

2 Main results and conjectures

2.1 Tori

We denote the n -cycle graph by C_n . This is a graph on n vertices containing a single cycle through all vertices.

Theorem 1. *There exist some absolute constants $c_1, c_2, \alpha, C_1, C_2 > 0$ such*

that for all $\lambda > 0$

$$\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[\lambda^2 \mathcal{S}(C_n) \leq C_1 \log^2 n] = 1, \quad (2.1)$$

$$\mathbb{P}_\lambda[\lambda^2 \mathcal{S}(C_n) < c_1 \log^2 n] \leq e^{-c_2 n^\alpha}. \quad (2.2)$$

We denote the d -dimensional torus of side length n by $\mathbb{T}_d(n)$. This is the Cayley graph of $(\mathbb{Z}/n\mathbb{Z})^d$ obtained by connecting each $x, y \in (\mathbb{Z}/n\mathbb{Z})^d$ which disagree only in one coordinate, by $\pm 1 \bmod n$.

Theorem 2. (i) *There exist $c, C > 0$ such that for all $\lambda > 0$*

$$\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[c \leq \lambda \mathcal{S}(\mathbb{T}_2(n)) / (\log n \log \log n) \leq C] = 1.$$

(ii) *For all $d > 2$ there exists $C'_d > 0$ such that for all $\lambda > 0$,*

$$\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[0.99d \leq \lambda \mathcal{S}(\mathbb{T}_d(n)) / \log n \leq C'_d] = 1.$$

2.2 Expanders

We denote by $\gamma(G)$ the **spectral-gap** of SRW on G , defined as the smallest non-zero eigenvalue of $I - P$, where P is the transition matrix of SRW on G . We say that a sequence of graphs G_n is an expander family if $\inf_n \gamma(G_n) > 0$. We say that G is a γ -**expander** if $\gamma(G) \geq \gamma$.

Theorem 3. *There exist some $\delta : \mathbb{N} \rightarrow (0, 1]$ which is $o(1)$ and an absolute constant $C > 0$ such that for every $\lambda \in (0, 1]$ and $\gamma \in (0, 2)$, there exists some $N_{\gamma, \lambda}$ so that for every connected regular γ -expander $G = (V, E)$ with $n := |V| \geq N_{\gamma, \lambda}$ we have that*

$$\mathbb{P}_\lambda[\lambda \mathcal{S}(G) > C\gamma^{-1} \log n] \leq \delta(n).$$

Remark 2.1. *Note that in the statement of Theorem 3, the (common) degree of the vertices plays no role. We think of γ and λ as being uniformly bounded away from 0, independently of the size of G , although our analysis remains valid even if this fails (as long as $\lambda\gamma$ does not tend to 0 too rapidly in terms of n).*

Let \mathcal{R}_t be the set of vertices which are visited by the process, when active particles vanishes after t steps. Consider a sequence of graphs $G_n := (V_n, E_n)$ with $|V_n| \rightarrow \infty$. Another natural question is whether for some fixed $t = t_\lambda$ we have that $\mathbb{P}_\lambda[|\mathcal{R}_t(G_n)| \geq \delta|V_n|] \geq \delta > 0$. While this problem is interesting by itself, a related problem will be central in proving an upper bound on \mathcal{S} in all the examples considered in this note. Consider the case that w_{plant} , the planted walker at \mathbf{o} , walks for $t = t_{|V|}$ steps (for some $t_{|V|}$ tending to infinity as $|V| \rightarrow \infty$), while the rest of the particles have lifetime M for some constant M . Denote the set of vertices which are visited by this modified process before it dies out by $\mathcal{R}_{t,M}$. In all of the examples analyzed in this paper we show that, $|\mathcal{R}_{t,M}| > \delta|V|$ w.h.p. for some $\delta > 0$, provided that $M = M_\lambda$ is sufficiently large and $t = t_{|V|} \rightarrow \infty$ as $|V| \rightarrow \infty$.

2.3 Conjectures

Definition 2.1. We say that a bijection $\varphi : V \rightarrow V$ is an automorphism of a graph $G = (V, E)$ if $\{u, v\} \in E$ iff $\{\varphi(u), \varphi(v)\} \in E$. A graph G is said to be vertex-transitive if the action of its automorphisms group, $\text{Aut}(G)$, on its vertices is transitive (i.e. $\{\varphi(v) : \varphi \in \text{Aut}(G)\} = V$ for all v).

Conjecture 2.1. There exists an absolute constants $C > 0$ such that for every finite connected vertex-transitive graph $G = (V, E)$ and all $\lambda, \delta > 0$

$$\mathbb{P}_\lambda[1 - \delta \leq \mathcal{S}(G)/t_\lambda(G) \leq C] \rightarrow 1, \quad \text{as } |V| \rightarrow \infty,$$

where $t_\lambda(G) := \min\{s : 2s/\kappa_s \geq \lambda^{-1} \log |V|\}$ and $\kappa_t := \min_v \sum_{i=0}^t p^i(v, v)$.

Remark 2.2. The fact that $\mathbb{P}_\lambda[\mathcal{S}(G) \geq (1 - \delta)t_\lambda(G)] \rightarrow 1$, as $|V| \rightarrow \infty$ follows from Theorem A. It is obtained by showing that w.h.p. there are many v 's such that even if at time 0 all of the vertices apart from v are activated, then v is not activated by time $(1 - \delta)t_\lambda(G)$.

The following conjecture is motivated by the results in [4] (see § 3.4 for more details).

Conjecture 2.2. There exist some $C_{d,\lambda}, \ell > 0$, such that for every sequence of finite connected graphs $G_n = (V_n, E_n)$ with $|V_n| \rightarrow \infty$ of maximal degree at most d , we have that $\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[\mathcal{S}(G_n) \leq C_{d,\lambda} \log^\ell |V_n|] = 1$.

We suspect that one can take above $\ell = 2$ and $C_{\lambda,d} = C\lambda^{-2}d^2$ for some absolute constant $C > 0$. Moreover, we suspect that for regular or vertex-transitive G , one can even take above, respectively, $C_{\lambda,d} = C\lambda^{-2}d$ or $C_{\lambda,d} = C\lambda^{-2}$ for some absolute constant $C > 0$ (c.f. [4, Conjectures 1.9 and 8.3]). See Examples 3.1-3.2 for more details about the dependence on d .

3 Propaedeutics

3.1 Notation

We denote $[k] := \{1, 2, \dots, k\}$ and $]k[:= \{0, 1, \dots, k\}$. We denote the cardinality of a set A by $|A|$. We write w.p. as a shorthand for “with probability”. We say that a sequence of events A_n defined, resp., with respect to some probabilistic model on a sequence of graphs $G_n := (V_n, E_n)$ with $|V_n| \rightarrow \infty$, holds w.h.p. (“with high probability”) if the probability of A_n tends to 1 as $n \rightarrow \infty$. We write $o(1)$ for terms which vanish as $n \rightarrow \infty$ (or some other index, which is clear from context). We write $f_n = o(g_n)$ if $f_n/g_n = o(1)$.

We shall use C, C', C_0, C_1, \dots (resp. c, c', c_0, c_1, \dots) to denote positive absolute constants which are sufficiently large (resp. small) to ensure that a certain inequality holds. Similarly, we use C_d, c_d or $C_{\lambda,d}, c_{\lambda,d}$ (etc.) to refer to positive constants, whose value depends on the parameters appearing in subscript (where the same convention applies regarding lower case and upper case letters). Different appearances of the same constant at different places may refer to different numeric values.

For SRW on a graph $G = (V, E)$, the **hitting time** of a set $A \subset V$ is $T_A := \inf\{t \geq 0 : X_t \in A\}$. Similarly, $T_A^+ := \inf\{t \geq 1 : X_t \in A\}$. When $A = \{x\}$ is a singleton, we instead write T_x and T_x^+ . Let P be the transition kernel of SRW on G . We denote by $p^t(u, v) := P^t(u, v)$ the t -steps transition probability from u to v . We denote by P_u the law of the entire walk, started from vertex u .

The distance $\text{dist}(x, y)$, between vertices x and y is the minimal amount of edges that one must pass in order to go from x to y . Vertices are said to be neighbors if they belong to a common edge.

3.2 A formal construction of the model

The following construction of the model is useful. We denote the set of $\text{Pois}(\lambda)$ (or $1 + \text{Pois}(\lambda)$ for the origin) frogs occupying vertex v at time 0 by $\mathcal{W}_v = \{w_1^v, \dots, w_{|\mathcal{W}_v|}^v\}$. We can assume that at time 0 there are infinitely many particles occupying each site v , $\mathcal{J}_v := \{w_i^v : i \in \mathbb{N}\}$ (where w_i^v is referred to as the i -th particle at v), but only the first $|\mathcal{W}_v|$ of them are actually involved in the dynamics of the model. We may think of each particle $w_i^v \in \mathcal{J}_v$ as first picking an infinite SRW $(S_t^{v,i})_{t \in \mathbb{Z}_+}$ according to P_v and only in the case that $i \leq |\mathcal{W}_v|$ and v is visited by some active particle, say at time s (for the first time), does w_i^v actually performs the first τ steps of the SRW it picked (i.e. its position in time $s + t$ is $S_t^{v,i}$ for all $t \in [\tau]$).

Let $G = (V, E)$ be a graph. A *walk* of length k in G is a sequence of $k + 1$ vertices (v_0, v_1, \dots, v_k) such that $\{v_i, v_{i+1}\} \in E$ for all $0 \leq i < k$. Let Γ_k be the collection of all walks of length k in G . We say that $w_i^v \in \mathcal{W}_v$ *picked* the path $\gamma = (\gamma_0, \dots, \gamma_k) \in \Gamma_k$ if $S_t^{v,i} = \gamma_t$ for all $t \in]k[$. For each $\gamma \in \Gamma_k$ let X_γ denote the number of particles in \mathcal{W}_{γ_0} , other than the particle planted at the origin, which picked the walk γ . For a walk $\gamma = (\gamma_0, \dots, \gamma_k) \in \Gamma_k$ for some $k \geq 1$, we denote $p(\gamma) := \prod_{i=0}^{k-1} P(\gamma_i, \gamma_{i+1})$. By Poisson thinning, we have that for every fixed k , the joint distribution of $(X_\gamma)_{\gamma \in \Gamma_k}$ (under \mathbb{P}_λ) is that of independent Poisson random variables with $\mathbb{E}_\lambda[X_\gamma] = \lambda p(\gamma)$ for all $\gamma \in \Gamma_k$.

Consider a collection of homogeneous Poisson processes on \mathbb{R}_+ with rate 1, $((M_v(s))_{s \geq 0})_{v \in \mathcal{V}}$ and a collection of simple random walks on \mathcal{G} , $\{(S_t^{v,i})_{t \in \mathbb{Z}_+}; v \in \mathcal{V}, i \in \mathbb{N}\}$, where for all i and x , $(S_t^{x,i})_{t \in \mathbb{Z}_+}$ is the walk picked by the i th frog in \mathcal{J}_x . We take the walks and the Poisson processes to be jointly independent. We take $|\mathcal{W}_u^\lambda| = M_u(\lambda) + 1_{u=\mathbf{o}}$, where \mathcal{W}_u^λ denotes the set of particles involved in the dynamics, whose initial position is u , when the density is taken to be λ . When clear from context, we omit the superscript λ and write \mathcal{W}_u . From this construction it is clear that the law of $\mathcal{S}(\mathcal{G})$ is stochastically decreasing in λ .

For $x, y \in \mathcal{V}(\mathcal{G})$ such that $x \neq y$ and $\tau \in \mathbb{N} \cup \{\infty\}$ let

$$\ell_\tau(x, y) := \inf\{j \leq \tau : S_j^{x,i} = y \text{ for some } i \leq |\mathcal{W}_x|\} \quad (3.1)$$

(where $\inf \emptyset = \infty$). The *activation time* of x (and also of \mathcal{W}_x) w.r.t. lifetime τ is

$$\text{AT}_\tau(x) := \inf\{\ell_\tau(x_0, x_1) + \dots + \ell_\tau(x_{m-1}, x_m)\}, \quad (3.2)$$

where the infimum is over all finite sequences $o = x_0, x_1, \dots, x_{m-1}, x_m = x$ where $x_i \in \mathcal{V}(\mathcal{G})$. The event $\text{AT}_\tau(x) = \infty$ is precisely the event that (for lifetime τ) site x is never visited by an active particle, while when finite, $\text{AT}_\tau(x)$ is the first time in which vertex x is visited by an active particle (for lifetime τ). The *susceptibility* of \mathcal{G} can be defined as

$$\mathcal{S}(\mathcal{G}) := \inf\{\tau : \max_{v \in \mathcal{V}} \text{AT}_\tau(v) < \infty\}. \quad (3.3)$$

3.3 A survey of some results from [12]

In this section we survey some result from [12] concerning the frog model on finite graphs. Theorem A will be used to bound \mathcal{S} from below in all of the examples considered in this note.

We denote the transition kernel of SRW on the underlying graph $G = (V, E)$ (which is clear from context) by P . We denote by $p^t(u, v) := P^t(u, v)$ the t -steps transition probability from u to v .

Theorem A. *[[12] Theorem 3] For every finite regular simple graph $G = (V, E)$ and all $\lambda > 0$*

$$\mathbb{P}_\lambda[\lambda \mathcal{S}(G) \geq \log |V| - 4 \log \log |V|] \rightarrow 1, \quad \text{as } |V| \rightarrow \infty. \quad (3.4)$$

Moreover, for all $\lambda > 0$ and $\delta \in [0, 1)$ we have that

$$\mathbb{P}_\lambda[\mathcal{S}(G) \geq t_{\lambda, \delta}(G)] \rightarrow 1, \quad \text{as } |V| \rightarrow \infty, \quad (3.5)$$

where $t_{\lambda, \delta}(G) := \min\{s : \frac{2s\lambda}{\kappa_s} \geq (1 - \delta) \log |V|\}$ and $\kappa_t := \min_v \sum_{i=0}^t p^i(v, v)$.

Remark 3.1. *Theorem A seems to be especially useful when G is vertex-transitive (see Conjecture 2.1). The bound offered by (3.5) is sharp up to a constant factor in all of the cases considered in this paper.*

Remark 3.2. *The argument in the proof of Theorem A is borrowed from Theorems 2 and 4.4 in [4]. It is possible to extend Theorem A to the case that G is non-regular. Adapting the argument from Theorem 4.3 in [4] gives*

$$\forall \alpha \in (0, 1), \exists c_\alpha > 0, \quad \mathbb{P}_\lambda[\lambda \mathcal{S}(G) \leq c_\alpha \log |V|] \leq \exp[-c'_\lambda r_*^2 |V|^\alpha], \quad (3.6)$$

where $r_* := \min_{u, v \in V} \frac{\deg(u)}{\deg(v)}$. This is meaningful as long as $r_* \geq |V|^{\beta-0.5}$ for some $\beta > 0$. We do not prove (3.6) (the details involved in the translation of (1.2) in [4] to (3.6) are similar to the ones involved in the translation of Theorems 2 and 4.4 in [4] to Theorem A above).

Remark 3.3. We note that for every regular graph $G = (V, E)$ we have that $t_{\lambda,0}(G) \leq C\lambda^{-2}\log^2|V|$ (e.g. [6, Lemma 2.4]), which is tight up to a constant factor as can be seen by considering the n -cycle, C_n , for which $t_{\lambda,1/2}(C_n) \geq c\lambda^{-2}\log^2|V|$.

Let $\mathcal{T}_{d,n}$ be the d -ary tree of depth n . The main result in [12] is

Theorem B. *There exist some $C_d, c_d > 0$ such that for all $\lambda > 0$ and $d \geq 2$*

$$\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[c_d \leq \lambda \mathcal{S}(\mathcal{T}_{d,n}) / (n \log n) \leq C_d] = 1.$$

Theorem C ([12] Proposition 1.1). *Let K_n be the complete graph on n vertices. For all $\lambda > 0$*

$$\forall \varepsilon \in (0, 1), \quad \lim_{n \rightarrow \infty} \mathbb{P}_\lambda[(1 - \varepsilon) \log n \leq \lambda \mathcal{S}(K_n) \leq (1 + \varepsilon) \log n] = 1.$$

Remark 3.4. *In light of Theorem A and Proposition C, it follows that K_n is the regular graph with asymptotically the smallest \mathcal{S} .*

3.4 Related models

The $A + B \rightarrow 2B$ family of models (e.g. [13, 14]) are defined by the following role: there are type A and B particles occupying a graph G , say with densities $\lambda_A, \lambda_B > 0$. They perform independent SRW with holding probabilities $p_A \in [0, 1]$ and $p_B \in [0, 1]$ (depending on the type). When a type B particle collides with a type A particle, the latter transforms into a type B particle. The frog model can be considered as a particular case of the above dynamics in which the type A particles are immobile ($p_A = 1$). Note that even when $p_A < 1$, one may consider the case in which the B particles have lifetime t and initially, only the particles at some vertex \mathbf{o} are of type B (and we may plant a B particle at \mathbf{o}). One can then define the susceptibility in an analogous manner, where the corresponding quantity is the number of steps required so that all particles are transformed into part B particles before the process dies out. We strongly believe that all of the results presented in this paper can be transferred into parallel results about the case of $p_A < 1$. Moreover, we also believe that the corresponding versions of Conjectures 2.1-2.2 are true also in the case of $p_A < 1$.

In [4], the first and third authors study the following model for a social network, called the random walks social network model, or for short, the SN

model. Given a graph $G = (V, E)$, consider $\text{Poisson}(|V|)$ walkers performing independent lazy simple random walks on G simultaneously, where the initial position of each walker is chosen independently w.p. proportional to the degrees. When two walkers visit the same vertex at the same time they are declared to be acquainted. The social connectivity time, $\text{SC}(G)$, is defined as the first time in which there is a path of acquaintances between every pair of walkers. The main result in [4] is that when the maximal degree of G is d , then w.h.p.

$$c \log |V| \leq \text{SC}(G) \leq C_d \log^3 |V|. \quad (3.7)$$

Moreover, $\text{SC}(G)$ is determined up to a constant factor in the cases that G is a regular expander or a d -dimensional torus ($d \geq 1$) of side length n .

Note that Conjecture 2.2 is the analog of (3.7) for the frog model (obtained by replacing $\text{SC}(G)$ above by $\mathcal{S}(G)$). In many examples, $\mathbb{E}[\text{SC}(G)]$ and $\mathbb{E}[\mathcal{S}(G)]$ are of the same order (when λ is fixed), and several techniques from [4] can be applied successfully to the frog model. Namely, the same technique used in [4] to prove general lower bounds on $\text{SC}(G)$ are used in the proof of Theorem A. Moreover, the analysis of the two models on expanders and on d -dimensional tori ($d \geq 1$) are similar (in all of these cases $\mathcal{S}(G)$ and $\text{SC}(G)$ are w.h.p. of the same order).

We now present a couple of examples with a large \mathcal{S} , demonstrating that $\mathcal{S}(G)$ may be large when the maximal degree of G is large, even if it is regular.

Example 3.1. Let G_n be the graph obtained by attaching a distinct vertex to each site of the complete graph on n vertices. It is easy to see that w.h.p. $c \leq \frac{\lambda \mathcal{S}(G_n)}{n \log n} \leq C$ for all fixed $\lambda > 0$.

Example 3.2. In general, $\mathbb{E}_1[\mathcal{S}(G)]$ may grow linearly in d , even when G is d -regular, as the following example demonstrates. Fix some $2 \leq d, n$ such that $2d \leq n$. Let J_k be a graph obtained from the complete graph on k vertices by deleting a single edge. Consider $\lceil n/d \rceil$ disjoint copies of J_d : $I_0, \dots, I_{\lceil n/d \rceil - 1}$, where for all $0 \leq j < \lceil n/d \rceil$, I_j is connected to I_{j+1} ($j+1$ is defined mod $\lceil n/d \rceil$) by a single edge that connects two degree $d-1$ vertices. This can be done so that the obtained graph is d -regular. It is easy to see that the obtained graph satisfies $\mathbb{E}_1[\mathcal{S}(G)] \geq cd \log(\frac{n}{d})$. We believe that this graph is the d regular graphs of size $n \pm o(n)$ with asymptotically the largest $\mathbb{E}_1[\mathcal{S}]$.

3.5 Auxiliary results

Definition 3.1. Let $G = (V, E)$ be a graph and $Y = (Y_v)_{v \in V}$ be some Bernoulli r.v.'s. We say that Y is **k -dependent** if $(Y_v)_{v \in V_1}$ and $(Y_u)_{u \in V_2}$ are independent, for all $V_1, V_2 \subset V$ such that

$$\min_{v_1 \in V_1, v_2 \in V_2} \text{dist}(v_1, v_2) > k.$$

We say that the random graph $(V, \{\{u, v\} \in E : Y_u = 1 = Y_v\})$ is k -dependent (or a k -dependent percolation process on G) if Y is k -dependent. We refer to the previous random graph as the graph associated with the percolation process induced by Y .

The following result is due the Liggett, Schonmann and Stacey [16].

Theorem D. Let $G = (V, E)$ be a graph of maximal degree Δ . For every $\alpha \in (0, 1)$ and $k \in \mathbb{N}$ there exists some $q(\alpha, \Delta, k)$ such that if $Y = (Y_v)_{v \in V}$ is a family of k -dependent Bernoulli r.v.'s with $\inf_{v \in V} \mathbb{P}[Y_v = 1] \geq q(\alpha, \Delta, k)$, then the law of Y stochastically dominated that of i.i.d. Bernoulli(α) r.v.'s $(X_v)_{v \in V}$.

Definition 3.2. Let $G = (V, E)$ be some graph. Let $\alpha \in [0, 1]$. Let $(X_v)_{v \in V}$ be i.i.d. Bernoulli(α) r.v.'s. The random graph $(V, \{\{u, v\} \in E : X_u = 1 = X_v\})$ is called Bernoulli site percolation on G with parameter α .

The following lemma is standard (e.g. [17]).

Lemma 3.1. Let $d \geq 2$. Then there exists some $p_c(d)$ (which is non-increasing in d) so that if $p \in (p_c(d), 1)$ then for some $C(d, p), R(d, p)$ and $c(p)$ (which can be chosen so that $C(d, p), R(d, p)$ tend to 0 and $c(p)$ tends to 1 as $p \rightarrow 1$) such that w.h.p. the random graph obtained from Bernoulli site percolation with parameter p on $\mathbb{T}_d(n)$ contains a unique connected component GC of size at least $R(d, p) \log n$. Moreover, GC satisfies w.h.p. that in every box of side-length $L(d, p) := \lceil (C(d, p) \log n)^{1/2} \rceil$, there are at least $c(p)L^d$ vertices belonging to GC. Lastly, there exists a decreasing function $f_d : \mathbb{N} \rightarrow [0, 1]$ such that $f_d = o(1)$ and for every n and $U \subset \mathbb{T}_d(n)$, the probability that $U \cap \text{GC}$ is empty is at most $f_d(|U|)$.

The following Lemma is essentially taken from [3]. We present its proof for the sake of completeness.

Lemma 3.2. *Let $G = (V, E)$ be a connected d -regular n -vertex γ -expander. Let $A \subset V$ and $R > 0$. Let $s = s_R := \lceil \gamma^{-1} \log(2^7 R) \rceil$. Consider*

$$H_A = H_{A_R} := \{v \in V : \sup_{t \geq s} |p_L^t(v, A) - \frac{|A|}{|V|}| \geq 1/4\}.$$

Let π be the uniform distribution on V . Write $\pi(B) := \frac{|B|}{|V|}$. Then

$$\pi(H_A) \leq \frac{1}{R} \pi(A) (1 - \pi(A)).$$

Proof. Consider $f : V \rightarrow \mathbb{R}$ defined by $f(x) := P_L^s(1_A - \pi(A))(x) = p_L^s(x, A) - \pi(A)$ (the middle term involves standard Markov chain notation, where for a transition Matrix Q , one writes $Q^t f(x) = \mathbb{E}_x[f(Y_t)]$, where $(Y_k)_{k=0}^\infty$ is the corresponding Markov chain). For $g : V \rightarrow \mathbb{R}$, denote $\mathbb{E}_\pi g := \sum_v \pi(v)g(v)$, $\|g\|_2 := \sqrt{\mathbb{E}_\pi g^2}$ and $\text{Var}_\pi g := \|g - \mathbb{E}_\pi g\|_2$. By the Poincaré (spectral gap) inequality $\|f\|_2^2 = \text{Var}_\pi P_L^s 1_A \leq (1 - \gamma/2)^{2s} \text{Var}_\pi 1_A \leq (2^7 R)^{-1} \text{Var}_\pi 1_A$.

Consider $f_*(x) := \sup_{t \geq 0} |P_L^t f(x)| = \sup_{t \geq s} |p_L^t(x, A) - \pi(A)|$. By Starr Maximal inequality (e.g. [3, Theorem 2.3]),

$$\|f_*\|_2^2 \leq 8\|f\|_2^2 \leq (16R)^{-1} \text{Var}_\pi 1_A.$$

Finally, note that $H_A := \{x : f_*^2(x) \geq 1/16\} \subset \{x : f_*^2(x) \geq R\|f_*\|_2^2 / \text{Var}_\pi 1_A\}$, and so by Markov inequality $\pi(H_A) \leq \text{Var}_\pi 1_A / R = R^{-1} \pi(A) (1 - \pi(A))$. \square

Corollary 3.1. *Let $G = (V, E)$ be a connected regular n -vertex γ -expander. Let $\lambda \in (0, 1]$. Let $A \subset V$. Let $r = r_{L, \lambda} := \lceil 7\lambda^{-1} \gamma^{-1} \log 2^9 \rceil$. Assume that $|A| < n/4$. Let (X_t) be SRW on G . Consider*

$$G_A = G_{A_{L, \lambda}} := \{a \in A : \mathbb{E}_a[|\{X_t : t \in [r]\} \setminus A|] \geq \lceil 9\lambda^{-1} \log 2 \rceil\}$$

(this is the set of all $a \in A$ such that the expected number of vertices in $V \setminus A$ belonging to the range of a length r SRW starting from a is at least $\lceil 9\lambda^{-1} \log 2 \rceil$). Then, provided that n is sufficiently large, we have that

$$|G_A| \geq \frac{3}{4}|A|.$$

Proof. Let H_{A_4} be as in Lemma 3.2. Consider $B := A \setminus H_{A_4}$. By Lemma 3.2 to conclude the proof, it suffices to show that $B \subset G_A$. Observe that SRW can be coupled with LSRW starting from the same initial position so

that they follow the same trajectory, with the LSRW spending at each site a random number of steps (with a Geometric(1/2) distribution) before moving to the next site. Hence it suffices to show that

$$p_b := \mathbb{E}_b[|\{X_t^L : t \in [r]\} \setminus A|] \geq \lceil 9\lambda^{-1} \log 2 \rceil,$$

for all $b \in B$, where (X_t^L) is a LSRW on G (instead of a SRW as in the assertion of the corollary). Let $b \in B$. By the definition of H_{A^c} we have that $p_L^t(b, A^c) \geq 1/2$ for all $t \geq k := \lceil \gamma^{-1} \log(2^9) \rceil$ (where $A^c := V \setminus A$). This implies that

$$e_b := \mathbb{E}_b[|\{1 \leq t \leq r : X_t^L \in A^c\}|] \geq (r - k)/2 \geq 3\lambda^{-1}\gamma^{-1} \log 2^9.$$

Finally, observe that

$$p_b \geq \frac{e_b}{\max_{v \in A^c} \sum_{i=0}^{r-1} p_L^i(v, v)} \geq \frac{e_b}{2\gamma^{-1} + r/n} \geq \lceil 9\lambda^{-1} \log 2 \rceil,$$

provided that n is sufficiently large. \square

4 Proofs

4.1 Tori

4.1.1 The cycle - Proof of Theorem 1

In this section we consider the case that G is the n -cycle, $C_n = (V(C_n), E_n)$. We start with the proof of (2.2). We fix $t = t_n = c_1 \lambda^{-2} \log^2 n$ for some constant c_1 to be determined later. Let $U_t(v)$ be the number of particles, whose initial position is not v , such that the walk they picked (picked in the sense of § 3.2) visits vertex $v \in V(C_n)$ by time t .

By considering a maximal collection of vertices which are all of distance at least $t + 1$ from one another and exploiting independence, in order to prove (2.2), it suffices to show that $\mathbb{P}_\lambda[U_t(v) = 0] \geq n^{-(1-\delta)}$, for some $\delta > 0$. Denote $\nu_t := \sum_{i=0}^t p^i(v, v) = \Theta(\sqrt{t+1})$. By Poisson thinning, if v is of distance at least $t + 1$ from \mathbf{o} then $U_t(v)$ has a Poisson distribution with mean (by reversibility)

$$\mu_t := \lambda \sum_{u: u \neq v} \mathbb{P}_u[T_v \leq t] \leq \frac{\lambda}{\nu_t} \sum_{u: u \neq v} \sum_{j=1}^t \mathbb{P}_u[T_v = j] \sum_{i=0}^{2t-j} p^i(v, v)$$

$$\leq \frac{C\lambda}{\sqrt{t+1}} \sum_{u:u \neq v} \sum_{i=1}^{2t} p^i(u, v) \leq \frac{C\lambda(2t)}{\sqrt{t+1}}.$$

Thus if $t < (\frac{\lambda \log n}{8C})^2$ we get that the probability that $U_t(v) = 0$ is at least $e^{-\mu t} \geq n^{-1/4}$.

We now prove (2.1). We fix $s = s_n := C_1 \lambda^{-2} \log^2 n$, where C_1 shall be determined shortly. For a vertex v let R_v and L_v be the line segments of length $k_n := \lceil 6\lambda^{-1} \log n \rceil$ to the right and left, resp., of v (not including v). Let $A_v = A_v(\lambda, n)$ be the event that for both R_v and L_v , there is at least one particle whose starting position is in that set (resp.), which picked a walk that reaches v by time s_n . It is not hard to verify that if C_1 is taken to be sufficiently large, then for every $u \in R_v \cup L_v$, the probability that a SRW starting from u would reach v by time s_n is at least 0.4 (in fact, we could have replaced 0.4 by any fixed number smaller than 1). Thus by Poisson thinning we have that $\mathbb{P}_\lambda[A_v] \geq 1 - 2e^{0.4\lambda k_n}$ for all v . Thus by the choice of k_n , a union bound on all the vertices yields that w.h.p. the event $A := \bigcap_{v \in V(C_n)} A_v$ occurs. Observe that on the event A , provided that the set of vertices which are visited by the process before it dies out contains an interval of size k_n , then deterministically, every site is visited before the process dies out. To conclude the proof, we argue that w.h.p. the planted particle w_{plant} visits in its length s_n walk at least $\lambda^{-1} \sqrt{\log n}$ vertices (we could have taken here any function of n which is $o(\sqrt{s_n})$). Given that this is indeed the case, then w.h.p. in one of the sites visited by w_{plant} there is some particle which visited at least k_n sites in its length s_n walk. \square

4.1.2 Tori - Proof of Theorem 2

Proof: By Theorem A we only need to prove the upper bounds on $\mathcal{S}(\mathbb{T}_d(n))$. Indeed, $\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[\lambda \mathcal{S}(\mathbb{T}_d(n)) \geq 0.99d \log n] = 1$ follows from (3.5) while for $d = 2$ $\lim_{n \rightarrow \infty} \mathbb{P}_\lambda[\lambda \mathcal{S}(\mathbb{T}_2(n)) \geq c \log n \log \log n] = 1$ (for some $c > 0$) follows from (3.4) using the fact that the expected number of returns to some vertex in $\mathbb{T}_2(n)$ by time $c\lambda^{-1} \log n \log \log n$ is at least $c' \log(c\lambda^{-1} \log n \log \log n)$.

We think of the vertices as being labeled by the set $[0, n-1]^d \cap \mathbb{Z}^d$. By abuse of notation, we denote the vertex set of $\mathbb{T}_d(n)$ again by $\mathbb{T}_d(n)$. A **box** of side length r is a set of the form $\{(x_1, \dots, x_d) : \forall i, \exists j_i \in \{0, \dots, r-1\}, x_i \equiv v_i + j_i \pmod{n}\}$ for some $(v_1, \dots, v_d) \in [0, n-1]^d \cap \mathbb{Z}^d$. We define the ℓ_p distance ($p \geq 1$) between $x, y \in \mathbb{T}_d(n)$, $\|x - y\|_p$, as $\min \|x' - y'\|_p$, where

the minimum is taken over all pairs x', y' in \mathbb{Z}^d such that $x' \equiv x$ and $y' \equiv y \pmod n$ (co-ordinate-wise) and $\|\cdot\|_p$ is the usual ℓ_p norm on \mathbb{R}^d . The same convention is utilized when we consider a renormalized torus of the form $\mathbb{T}_d(\lfloor n/r \rfloor)$ (when we replace $\pmod n$ by $\pmod{\lfloor n/r \rfloor}$).

We split the particles into two sets. Let \mathcal{W}_v be the set of particles whose initial position is v . For each v we partition \mathcal{W}_v into two disjoint sets \mathcal{W}_v^1 and \mathcal{W}_v^2 so that $(|\mathcal{W}_v^i|)_{(v,i) \in \mathbb{T}_d(n) \times \{1,2\}}$ are mutually independent having a $\text{Pois}(\frac{\lambda}{2})$ distribution, apart from \mathcal{W}_o^1 , for which $|\mathcal{W}_o^1| - 1 \sim \text{Pois}(\frac{\lambda}{2})$ (i.e. the planted particle w_{plant} belongs to \mathcal{W}_o^1). We call the particles in $\cup_{v \in \mathcal{V}} \mathcal{W}_v^i$ *type i particles* ($i = 1, 2$). We first analyze the dynamics only w.r.t. the type 1 particles (as if there are no other type 2 particles). Given a certain lifetime, this dynamics is the same as the frog model with density $\lambda/2$ with that lifetime. However, below we shall take the lifetime of w_{plant} to be one tending to infinity as $n \rightarrow \infty$, while the rest of the type 1 particles shall have a constant lifetime. We refer to this setup as the *type 1 dynamics*. Denote

$$t_2 = t_2(n, \lambda) = M(2)\lambda^{-1} \log n \log \log n$$

$$t_d = t_d(n, \lambda) = M(d)\lambda^{-1} \log n,$$

where $M(d)$ ($d \geq 2$) are constants to be determined later. We consider the setup in which w_{plant} has lifetime t_d , while the rest of the type 1 particles have lifetime $s = s(d, \lambda)$ to be determined later.

The following proposition allows us to reduce the proof of Theorem 2 to an easier problem. Let D be the collection of all vertices which are visited by the type 1 dynamics before it dies out. By Proposition 4.1 it suffices to show that D satisfies the following spatial condition w.h.p. (provided that $s(d, \lambda)$ is sufficiently large). There exist some $C(d), c(d) > 0$ such that for every box B of side length $L := \lceil (C(d) \max(\lambda^{-1}, 1) \log n)^{1/2} \rceil$ we have that $|D \cap B| \geq c(d)|B|$. To see this, take A in the proposition to be D and take the particles to be the type 2 particles.

Proposition 4.1. *Let $d \geq 2$ and $\lambda > 0$. Fix some $C(d)$ and let $n \in \mathbb{N}$ be such that $L := \lceil (C(d) \max(\lambda^{-1}, 1) \log n)^{1/2} \rceil \leq n^{1/2}$. Let $A \subset \mathbb{T}_d(n)$ be such that for every box B of side length L we have that $|A \cap B| \geq c(d)|B|$. Assume that at each $a \in A$ there are initially $\text{Pois}(\lambda/2)$ (independently for different vertices) particles performing simultaneously independent SRWs of length t_d . Let D denote the union of the ranges of these walks. Then provided that $M(d)$ is taken to be sufficiently large, $\mathbb{P}[D \neq \mathbb{T}_d(n)] \leq C'n^{-d}$.*

Proof. Let $p_v := \sum_{a \in A} P_a[T_v \leq t_d]$ and $e_v := \sum_{a \in A} \sum_{i=0}^{t_d} P^i(a, v)$. Let $\alpha_d := \sum_{i=0}^{t_d} P^i(\mathbf{o}, \mathbf{o})$. Note that $\alpha_d \leq C_2$ for all $d \geq 3$ and $\alpha_2 \leq C_3 \log \log n$, provided that n is sufficiently large. By Lemma 4.1 we can pick $M(d)$ so that $e_v/\alpha_d > 4d\lambda^{-1} \log n$. However, $p_v \geq e_v/\alpha_d$. Finally, by Poisson thinning, the number of particles (from A) which visit v by time t_d has a Poisson distribution with mean $(\lambda/2)p_v$. The proof is concluded by a union bound over all $v \in V$. \square

Lemma 4.1. *Let $d \geq 2$. Fix some $C(d)$ and let $n \in \mathbb{N}$ be such that $L := \lceil (C(d) \log n)^{1/2} \rceil \leq n^{1/2}$. Let $B \subset \mathbb{T}_d(n)$ be a box of side length L . Let $x \in B$. Then there exists some $c(d) > 0$ such that for all $t \geq (9d \log d)L^2$ and all $y \in \mathbb{T}_d(n)$ such that $\|x - y\|_2 \leq 3\sqrt{t}$ we have that $P^t(x, y) + P^{t+1}(x, y) \geq c \frac{1}{|B|} (P^t(y, B) + P^{t+1}(y, B))$. In particular, if $A \subset B$ is of size at least $c'(d)|B|$, then for some $\epsilon(d) > 0$ (independent of $C(d)$) we have that*

$$\sum_{a \in A} \sum_{t=0}^{(18d \log d)L^2} P^t(a, y) \geq \epsilon(d)L^2.$$

The assertion of the lemma is a consequence of the local CLT and reversibility, c.f. [4, Lemma 7.2].

As we now explain, our strategy for verifying the spatial condition described in the paragraph preceding Proposition 4.1 is to reduce it to an even simpler condition involving some 2-dependent (see Definition 3.1) auxiliary percolation process on a renormalized torus, whose sites are boxes of the original torus. In order to study the type 1 dynamics, we split the type 1 particles into two sets. For each v we partition \mathcal{W}_v^1 into two disjoint sets \mathcal{W}_v^a and \mathcal{W}_v^b so that $(|\mathcal{W}_v^i|)_{(v,i) \in \mathbb{T}_d(n) \times \{a,b\}}$ are mutually independent having a $\text{Pois}(\frac{\lambda}{4})$ distribution, apart from $\mathcal{W}_\mathbf{o}^a$, for which $|\mathcal{W}_\mathbf{o}^a| - 1 \sim \text{Pois}(\frac{\lambda}{4})$ (i.e. the planted particle belongs to $\mathcal{W}_\mathbf{o}^a$). We call the particles in $\cup_{v \in V} \mathcal{W}_v^i$ *type i particles* ($i = a, b$). As before, the lifetime of all these particles is taken to be $s(d, \lambda)$ apart from w_{plant} whose lifetime is t_d . The type a and type b dynamics are defined by ignoring the other type of particles. For a collection of particles \mathcal{W}' we denote by $\mathcal{R}_k(\mathcal{W}')$ the union of the ranges of the first k steps of the walks picked by the particles in \mathcal{W}' .

Fix some integer $r = r(d, \lambda)$ to be determined later. Consider a partition of $\mathbb{T}_d(n)$ into $(\lfloor n/r \rfloor)^d$ boxes of side length r (apart from $O((n/r)^{d-1})$ boxes which may be of uneven side lengths between r and $2r$). The boxes inherit

naturally the structure of $\mathbb{T}_d(\lfloor n/r \rfloor)$. Namely, for every $v \in \mathbb{T}_d(\lfloor n/r \rfloor)$ we denote by B_v the unique box in the partition containing rv .

Definition 4.1. Let $v \in \mathbb{T}_d(\lfloor n/r \rfloor)$. We define the B_v dynamics, started at $x \in B_v$ to be the variation of type a dynamics with lifetime $s = s(d, \lambda)$ (where initially only \mathcal{W}_x^a (type a particles at x) are activated; here there is no planted particle at x) in which $\mathcal{W}_y^a \sim \text{Pois}(\lambda/4)$ for all $y \in B_v$ and initially there are no particles outside B_v . We refer to the aforementioned dynamics as the B_v dynamics started from x . Let A_x be the collection of vertices in B_v that are visited by the B_v dynamics started from x , before it dies out.

Definition 4.2. We say that $x \in B_v$ is **good** if $|A_x| \geq \frac{1}{4}|B_v|$. If there is at least one good site in B_v , we denote the one which is closest to the center of B_v (breaking ties, say using the lexicographic order) by \mathbf{o}_v . We say that B_v is **good** if (1) B_u contains at least one good vertex for all u such that $\|u - v\|_1 \leq 1$ and (2) $\mathbf{o}_u \in \mathcal{R}_s(\cup_{y \in A_{\mathbf{o}_v}} \mathcal{W}_y^b)$ for all u such that $\|u - v\|_1 = 1$. Finally, for all $v \in \mathbb{T}_d(\lfloor n/r \rfloor)$ we set Y_v to be the indicator of the event that B_v is good.

Let $H = (\lfloor n/r \rfloor^d, \tilde{E})$ be the random subgraph of $\mathbb{T}_d(\lfloor n/r \rfloor)$ obtained by setting $\tilde{E} := \{\{u, v\} : u, v \in \lfloor n/r \rfloor^d, \|u - v\|_1 = 1, Y_v = 1 = Y_u\}$. This is the graph associated with the percolation process (on $\mathbb{T}_d(\lfloor n/r \rfloor)$) induced by $(Y_v)_{v \in \mathbb{T}_d(\lfloor n/r \rfloor)}$ (see Definition 3.1). Observe that by construction, if we look at the type 1 dynamics started at \mathbf{o}_v for some v satisfying $Y_v = 1$ (in which all particles have lifetime s and there is no planted particle at \mathbf{o}_v) then by construction, the collection of vertices which are activated before the process dies out, contains

$$\cup_{u: u \text{ is in the connected component of } v \text{ in } H} A_{\mathbf{o}_u}.$$

We will show that if r and s are chosen to be sufficiently large, then $\mathbb{P}_\lambda[Y_v = 1]$ can be made arbitrarily close to 1 (this probability is independent of v). We now explain how this implies that D (the collection of sites visited by the type 1 dynamics) satisfies the desired spatial property from the paragraph preceding Proposition 4.1:

Note that $Y := (Y_v)_{v \in \mathbb{T}_d(\lfloor n/r \rfloor)}$ is 2-dependent. For every $p \in (0, 1)$, by setting s and r so that $\mathbb{P}_\lambda[Y_v = 1]$ is sufficiently close to 1, using Theorem D we get that Y stochastically dominates i.i.d. Bernoulli p random variables $(X_v)_{v \in \mathbb{T}_d(\lfloor n/r \rfloor)}$. Let GC be the largest connected component H (the graph

associated with the percolation induced by Y). By Lemma 3.1 it follows that

$$A := \cup_{v \in \text{GC}} A_{\mathbf{o}_v}$$

satisfies the spatial condition from the paragraph preceding Proposition 4.1 w.h.p.. Hence, by the paragraph following Definition 4.2 it suffices to show that w.h.p. there is some $v \in \text{GC}$ such that $\mathbf{o}_v \in \mathcal{R}_{t_d}(w_{\text{plant}})$.

Again, using Theorem D it is not hard to show that w.h.p.

$$F := \{v \in \mathbb{T}_d(\lfloor n/r \rfloor) : Y_v = 1 \text{ and } \mathbf{o}_v \in \mathcal{R}_{t_d}(\{w_{\text{plant}}\})\}$$

is of size at least $t_d^{1/3}$ (in fact we can replace $t_d^{1/3}$ by $f(t_d)$ any $f(n) = o(n)$). We leave this as an exercise. Using Lemma 3.1 again, it follows that $F \cap \text{GC}$ is non-empty w.h.p.. On this event, by construction we have that D (the collection of sites visited by the type 1 process) contains A .

To conclude the proof, we now verify that if r and s are chosen to be sufficiently large, then $\mathbb{P}_\lambda[Y_v = 1]$ can be made arbitrarily close to 1. For $d = 2$ this is an easy consequence of the fact that SRW on \mathbb{Z}^2 is recurrent. We leave the details as an exercise. We now consider the case $d \geq 3$. We take r so that $r\lambda \geq C_1(d)$ for some $C_1(d)$ to be determined later. Set $s := dr^2$.

Fix some $v \in \mathbb{T}_d(\lfloor n/r \rfloor)$. We will show that the probability that B_v contains at least one good site can be made arbitrarily close to 1 by picking $C_1(d)$ appropriately. Before doing so, we first explain how this implies the same for the event $Y_v = 1$. Indeed, by a union bound over the neighbors of v also the probability that B_u contains at least one good site for all u such that $\|u - v\|_1 \leq 1$ (this is the first requirement in the definition of a good box) can be made arbitrarily close to 1. We now show that, given that the first requirement from the definition of a good box is satisfied for B_v , then also the second requirement is satisfied with (conditional) probability that can be made arbitrary close to 1 (again using a union bound over the neighbors of v) by taking $C_1(d)$ to be sufficiently large. Let u be such that $\|u - v\|_1 \leq 1$ and let $u' \in B_u$. Observe that given that $A_{\mathbf{o}_v} = A'$ for some $A' \subset B_v$ of size at least $\frac{1}{4}|B_v| \geq \frac{r^d}{4}$ and that $\mathbf{o}_u = u'$, the number of particles $w \in \cup_{y \in A_{\mathbf{o}_v}} \mathcal{W}_y^b$ satisfying that $\mathbf{o}_u \in \mathcal{R}_s(\{w\})$, has (by Poisson thinning) a Poisson distribution whose mean is at least $\lambda c_1(d)r^2$, where this estimate holds uniformly in A' and u' as above. To see this, recall that for $d > 2$ the Green's function $G(x, y) := \sum_{i \geq 0} p^i(x, y)$ is proportional to $\|x - y\|_2^{2-d}$. Moreover, for all $k \geq \|x - y\|_2^2$, we have that $G(x, y)$ is also proportional

to both $G_k(x, y) := \sum_{i \leq k} p^i(x, y)$ (by the local CLT) and $\mathbb{P}_x[T_y < k]$ (by transience). This implies that for all $v' \in B_v$ and $u' \in B_u$ we have that $\mathbb{P}_{v'}[T_{u'} < s] \geq c_2(d)r^{-(d-2)}$. Using this, it is easy to obtain the aforementioned $\lambda c_1(d)r^2$ lower bound on the mean.

To conclude the proof, we now verify that the probability that B_v contains at least one good site can be made arbitrarily close to 1 by picking $C_1(d)$ appropriately. Let $x \in B_v$. The argument below shows that given that \mathcal{W}_x^a is non-empty and $|\mathcal{R}_s(\mathcal{W}_x^a)| \geq 2r$, we get that x is good with probability that can be made arbitrarily close to 1, by taking $C_1(d)$ to be sufficiently large.

We may expose A_x by first exposing $\mathcal{R}_s(\mathcal{W}_x^a)$ and then continue by exposing $\mathcal{R}_s(\mathcal{W}_{x_1}^a)$ for some $x_1 \in \mathcal{R}_s(\mathcal{W}_x^a)$. Continue in this fashion, by exposing in the i th stage $\mathcal{R}_s(\mathcal{W}_{x_i}^a)$ for some vertex $x_i \in \cup_{j=0}^{i-1} \mathcal{R}_s(\mathcal{W}_{x_j}^a) \cap B_v \setminus \{x_j : 0 \leq j \leq i-1\}$ (with the convention that $x_0 := x$), where x_i is chosen according to some predetermined rule. Denote

$$\mathcal{B}_i := \cup_{j=0}^i \mathcal{R}_s(\mathcal{W}_{x_j}^a) \cap B_v.$$

Observe that as long as $|\mathcal{B}_i| > i+1$ we can pick some x_{i+1} and continue the above exploration process. Let U_i be the event that $i+1 < |\mathcal{B}_i| < \frac{1}{4}|B_v|$. We argue that there exists $c_3(d) > 0$ (independent of $C_1(d)$) so that a.s.

$$\mathbb{E}[|\mathcal{B}_{i+1} \setminus \mathcal{B}_i| 1_{U_i} \mid \mathcal{R}_s(\mathcal{W}_{x_0}^a), \dots, \mathcal{R}_s(\mathcal{W}_{x_i}^a)] \geq c_3(d)\lambda r^2 1_{U_i}. \quad (4.1)$$

Indeed, on U_i we have that $|B_v \setminus \mathcal{B}_i| > \frac{3}{4}|B_v|$ and thus (4.1) follows by the aforementioned discussion regarding the Green's function. Let Z_i be the indicator of the event that either U_i^c occurs, or that $|\mathcal{B}_{i+1} \setminus \mathcal{B}_i| > \frac{1}{2}c_3(d)r^2$. Using Paley-Zygmund inequality and (4.1), we obtain that the joint law of Z_1, \dots, Z_{r^d} stochastically dominates that of i.i.d. Bernoulli r.v.'s with mean $c_4(d) > 0$. Hence, by Chernoff bound, the probability that $\sum_{i=1}^j Z_i > \frac{1}{2}c_4(d)j - r$ for all $1 \leq j \leq |B_v|$ can be made arbitrarily close to 1, by setting $C_1(d)$ to be sufficiently large. Note that if this occurs, and in addition also $|\mathcal{B}_0| > 2r$, then as long as $C_1(d)$ is sufficiently large, we have that deterministically $|A_x| > \frac{3}{4}|B_v|$. The proof is concluded by noting that for every $x \in B_v$, the probability that $\mathcal{B}_0 = \mathcal{R}_s(\mathcal{W}_x^a)$ is of size at least $2r$ is at least $c_5(d)\lambda$ and that this occurs independently for all $x \in B_v$. Again, as $r^d\lambda > C_1(d)$, the probability that there is no $x \in B_v$ such that $|\mathcal{R}_s(\mathcal{W}_x^a)| > 2r$ can be made arbitrarily small by taking $C_1(d)$ to be sufficiently large. \square

4.2 Expanders - Proof of Theorem 3

In this section we study the case that G is a d -regular expander. It is not difficult to extend the results to the case G is an expander of maximal degree d . We note that the arguments presented in this section are inspired by techniques from [4, Theorem 5] and [8, Theorem 3]. However, the analysis below includes some new ideas. In particular, the usage of a Maximal inequality in the proof of Theorem 4.1 (Lemma 3.2) is novel.

Below we consider also lazy SRW (LSRW) on G , defined as follows. In each step, the walk either stays put w.p. $1/2$ or w.p. $1/2$ moves to one of the neighbors of the current position of the walk, chosen from the uniform distribution over the neighbor set. Clearly, the frog model evolves slower in the case that some of the particles perform LSRW instead of SRW.

Note that the transition matrix of LSRW on G has transition matrix $P_L := \frac{1}{2}(I+P)$ whose spectral gap is half the spectral gap of P . We denote the t -step transition probabilities with respect to LSRW by $p_L^t(u, v) := P_L^t(u, v)$. LSRW on a regular γ -expander $G = (V, E)$ mixes rapidly in the following sense

$$\left| \max_{x, y \in V} p_L^t(x, y) - |V|^{-1} \right| \leq (1 - \gamma/2)^t, \quad \text{for all } t. \quad (4.2)$$

Indeed $\max_{x \in V} |p_L^t(x, x) - |V|^{-1}| \leq (1 - \gamma/2)^t$ follows from the spectral decomposition of $p_L^t(x, x)$ along with the non-negativity of the eigenvalues of P_L . However, for LSRW on a regular graph $\max_{x, y \in V} |p_L^t(x, y) - |V|^{-1}| = \max_{x \in V} |p_L^t(x, x) - |V|^{-1}|$ (e.g. [7, (2.2)] for even t and [15, p. 135] for moving from even t to odd t using laziness).

Consider the case that w_{plant} , the planted walker at \mathbf{o} , walks for $t = t_{|V|}$ steps, while the rest of the particles have lifetime M for some constant M . Recall that the set of vertices which are visited by this modified process before it dies out is denoted by $\mathcal{R}_{t, M}$.

Theorem 4.1. *There exist some $\delta(n) = o(1)$ and an absolute constant $M > 0$ and some $N_{\lambda, \gamma}$, so that for all λ, γ and all regular γ -expander of size $n \geq N_{\lambda, \gamma}$, G we have that*

$$\mathbb{P}_\lambda[|\mathcal{R}_{\lceil M\lambda^{-1}\gamma^{-1} \log n \rceil, \lceil M\lambda^{-1}\gamma^{-1} \rceil}(G)| < n/4] \leq \delta(n),$$

Remark 4.1. *We could have replaced the $\log n$ term above by any function of n which tends to infinity as $n \rightarrow \infty$.*

We argue that Theorem 3 follows from Theorem 4.1. Indeed, we can partition the particles into two independent sets, each with density $\lambda/2$, with w_{plant} belonging to the first set of particles. Call the first (resp. second) set type 1 (resp. 2) particles. We can apply Theorem 4.1 to the dynamics associated with the type 1 particles (as if there are no type 2 particles; This dynamics is precisely the frog model with parameter $\lambda/2$). Denote by A the collection of vertices visited by the type 1 dynamics before it dies out. By Corollary 4.1, given that $|A| \geq n/4$, we have that w.h.p. V equals to the union of the ranges of the walks of length $\lceil C\lambda^{-1}\gamma^{-1}\log n \rceil$ performed by the type 2 particles initially occupying A . \square

Lemma 4.2. *Let $G = (V, E)$ be a connected regular n -vertex γ -expander. Then there exists $C > 0$ such that for all $v, u \in V$ the probability that a LSRW started at v would visit u by time $t := \lceil C\gamma^{-1}\log n \rceil$ is at least $\frac{1}{16} \min(\frac{\gamma t}{n}, 1)$.*

Proof. Let $(X_s)_{s \geq 0}$ be a LSRW on G started at v . We may assume $u \neq v$. Let $Y_i = 1_{X_i=u}$ and $Y := \sum_{i=1}^t Y_i$, where $t := \lceil C\gamma^{-1}\log n \rceil$, for some constant C , to be determined shortly. By (4.2), if C is sufficiently large, $\mathbb{E}[Y] \geq t/(2n)$, whereas $\mathbb{E}[Y_i Y_{i+j}] \leq \mathbb{E}[Y_i] p_L^j(u, u) \leq \mathbb{E}[Y_i](n^{-1} + (1 - \gamma/2)^j)$. Summing over $j \in [t]$ and then over $i \in [t]$ gives $\mathbb{E}[Y^2] \leq 2\mathbb{E}[Y](2\gamma^{-1} + \frac{t}{n})$. Finally,

$$\mathbb{P}[Y > 0] \geq \frac{(\mathbb{E}[Y])^2}{\mathbb{E}[Y^2]} \geq \frac{\mathbb{E}[Y]}{2(2\gamma^{-1} + \frac{t}{n})} \geq \frac{t/(2n)}{4 \max(2\gamma^{-1}, \frac{t}{n})} \geq \frac{1}{16} \min(\frac{\gamma t}{n}, 1).$$

\square

The following corollary is an immediate consequence of 4.2, obtained by a union bound over V , using Poisson thinning and independence.

Corollary 4.1. *Let $G = (V, E)$ be a connected regular n -vertex γ -expander. Let $\ell := \lceil n/4 \rceil$. Let $(v_i)_{i=0}^\ell$ be an arbitrary collection of distinct vertices. Assume that at each of these vertices there are initially $\text{Pois}(\lambda)$ particles, independently, and that the particles perform simultaneously independent LSRW on G . Let A_s be the event that the union of the first s steps performed by the particles equals V . Then there exist some constants C, n' (independent of G, d, γ and the sequence $(v_i)_{i=0}^\ell$) such that $\mathbb{P}[A_t] \geq 1 - n^{-2}$, for $t := \lceil C\lambda^{-1}\gamma^{-1}\log n \rceil$, as long as $n \geq n'$.*

Proof of Theorem 4.1. We use an exploration process due to Benjamini, Nachmias and Peres [5]. Let γ be the spectral gap of SRW on $G = (V, E)$. We

set

$$t = t_{C,\lambda} := \lceil \frac{C}{\lambda\gamma} \rceil,$$

where C shall be determined later. For a collection of particles, \mathcal{W}' let $\mathcal{R}_\ell(\mathcal{W}')$ denote the union of the ranges of the length ℓ walks performed by the particles in \mathcal{W}' (if \mathcal{W}' is empty, then this union is defined to be the empty set). Let \mathcal{W}_v be the particles whose initial position is v . For every $v \in V \setminus \{\mathbf{o}\}$ if \mathcal{W}_v is non-empty pick some $w_v \in \mathcal{W}_v$ and set $\mathcal{W}'_v := \{w_v\}$. Otherwise set \mathcal{W}'_v to be the empty set. For $v = \mathbf{o}$ we set $\mathcal{W}'_{\mathbf{o}}$ to be the empty set if $\mathcal{W}_{\mathbf{o}} \setminus \{w_{\text{plant}}\}$ is empty and otherwise, we set $\mathcal{W}'_{\mathbf{o}} := \{w_{\mathbf{o}}\}$ for some $w_{\mathbf{o}} \in \mathcal{W}_{\mathbf{o}} \setminus \{w_{\text{plant}}\}$. Below instead of working with the sets \mathcal{W}_v , we shall work with the sets \mathcal{W}'_v ($v \in V$). This is done in order to apply Azuma inequality at the end of the proof (otherwise, we would not have a good bound on the increments of the submartingale).

For every set A let

$$G_A := \{a \in A : \mathbb{E}_a[|\{X_s : s \in [t]\} \setminus A|] \geq \lceil 9\lambda^{-1} \log 2 \rceil\},$$

be the collection of all vertices $a \in A$ such that the expected number of vertices in $V \setminus A$ which are visited by a length t SRW started from a is at least $K_\lambda := \lceil 9\lambda^{-1} \log 2 \rceil$. The exact value of K_λ is unimportant. Below we shall only use the fact that $\mathbb{P}[\text{Pois}(\lambda) \neq 0]K_\lambda = (1 - e^{-\lambda})K_\lambda \geq 2$, for all $\lambda \in (0, 1]$. By Corollary 3.1 we can pick C so that for every A of size at most $n/4$ we have that $|G_A| \geq \frac{3}{4}|A|$.

First expose $\mathcal{R}_{\lceil M\lambda^{-1}\gamma^{-1} \log n \rceil}(\mathbf{o})$ and set $\mathcal{A}_0 := \mathcal{R}_{\lceil M\lambda^{-1}\gamma^{-1} \log n \rceil}(\mathbf{o}) \setminus \{\mathbf{o}\}$ and $\mathcal{U}_0 = \emptyset$. Assume that we have already defined $\mathcal{A}_0, \mathcal{A}_1, \dots, \mathcal{A}_j$ and $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_j$ in the following manner.

- (1) For all $0 \leq i \leq j$ we have that $\mathcal{U}_i := \{v_1, v_2, \dots, v_i\}$ and that

$$\mathcal{B}_i := \mathcal{A}_i \cup \mathcal{U}_i = \mathcal{A}_0 \cup (\cup_{k=1}^i \mathcal{R}_t(\mathcal{W}'_{v_k})).$$

Moreover, assume that $|\mathcal{B}_j| < n/4$.

- (2) For all $1 \leq i < j$ we have that $|\mathcal{B}_i| > \frac{4}{3}i$ and thus (using $|G_{\mathcal{B}_i}| \geq \frac{3}{4}|\mathcal{B}_i| > i$) it must be the case that $G_{\mathcal{B}_i} \cap \mathcal{A}_i$ is non-empty.
- (3) At stage $1 \leq i \leq j$ of the exploration process we expose $\mathcal{R}_t(\mathcal{W}'_{v_i})$ for some vertex $v_i \in \mathcal{A}_{i-1} \cap G_{\mathcal{B}_{i-1}}$. We then set $\mathcal{U}_i = \mathcal{U}_{i-1} \cup \{v_i\}$ and $\mathcal{A}_i = \mathcal{A}_{i-1} \cup \mathcal{R}_t(\mathcal{W}'_{v_i}) \setminus \mathcal{U}_i$.

The exploration process is terminated once $|\mathcal{B}_j| \geq n/4$ or $|\mathcal{B}_j| \leq \frac{4}{3}j$. Let J be the stage in which the exploration process is terminated. Let \mathcal{F}_i be the natural filtration of the exploration process by stage i , including. Let U_i be the event that $J > i$ and that $|\mathcal{W}'_{v_{i+1}}| > 0$. By the definition of $G_{\mathcal{B}_i}$, it follows that a.s.

$$\mathbb{E}[|\mathcal{B}_{i+1} \setminus \mathcal{B}_i| 1_{U_i} \mid \mathcal{F}_i, |\mathcal{W}'_{v_{i+1}}|] \geq 1_{U_i} [9\lambda^{-1} \log 2] > \frac{2}{(1 - e^{-\lambda})} \cdot 1_{U_i}.$$

Hence a.s.

$$\mathbb{E}[|\mathcal{B}_{i+1} \setminus \mathcal{B}_i| 1_{J>i} \mid \mathcal{F}_i] \geq 1_{J>i} \mathbb{P}[|\mathcal{W}'_{v_{i+1}}| > 0 \mid J > i] [9\lambda^{-1} \log 2] \geq 2 \cdot 1_{J>i}.$$

It follows that $\mathcal{M}_i := |\mathcal{B}_{i \wedge J}| - 2(i \wedge J)$ (where $i \wedge J := \min(i, J)$) is a submartingale with respect to the filtration (\mathcal{F}_i) . Observe that $|\mathcal{B}_1| \geq \lambda^{-1} \sqrt{\log n}$ w.h.p. (we could have replaced $\sqrt{\log n}$ by any $o(\log n)$ term; We omit the details). Using (this in conjunction with) Azuma inequality (after first conditioning on $|\mathcal{B}_1| \geq \lambda^{-1} \sqrt{\log n}$), it is not hard to verify that w.h.p. we have that $\mathcal{M}_i > -\frac{2}{3}i$ for all i (c.f. [5]). The last event implies that for all $i < J$ we have that $|\mathcal{B}_{i \wedge J}| > \frac{4}{3}i$, which means that $|\mathcal{B}_J| \geq n/4$, as desired. \square

Acknowledgements

.

References

- [1] OSM Alves, F.P. Machado, and S.Y. Popov. Phase transition for the frog model1. *Electronic Journal of probability*, 7(16):1–21, 2002.
- [2] OSM Alves, FP Machado, and S.Y. Popov. The shape theorem for the frog model. *The Annals of Applied Probability*, 12(2):533–546, 2002.
- [3] Ridhhipratim Basu, Jonathan Hermon, and Yuval Peres. Characterization of cutoff for reversible markov chains. *Ann. Probab., to appear*. <http://arxiv.org/abs/1409.3250>, 2013.
- [4] Itai Benjamini and Jonathan Hermon. Rapid social connectivity. *arXiv preprint arXiv:1608.07621*, 2016.

- [5] Itai Benjamini, Asaf Nachmias, and Yuval Peres. Is the critical percolation probability local? *Probability theory and related fields*, 149(1-2):261–269, 2011.
- [6] Lucas Boczkowski, Yuval Peres, and Perla Sousi. Sensitivity of mixing times in eulerian digraphs. *arXiv preprint arXiv:1603.05639*, 2016.
- [7] Sharad Goel, Ravi Montenegro, and Prasad Tetali. Mixing time bounds via the spectral profile. *Electron. J. Probab*, 11(1):1–26, 2006.
- [8] Jonathan Hermon, Ben Morris, Chuan Qin, and Allan Sly. The social network model on infinite graphs. *Arxiv preprint*, 2016.
- [9] Christopher Hoffman, Tobias Johnson, and Matthew Junge. Recurrence and transience for the frog model on trees. *arXiv preprint arXiv:1404.6238*, 2014.
- [10] Christopher Hoffman, Tobias Johnson, and Matthew Junge. From transience to recurrence with poisson tree frogs. *arXiv preprint arXiv:1501.05874*, 2015.
- [11] Tobias Johnson and Matthew Junge. The critical density for the frog model is the degree of the tree. *arXiv preprint arXiv:1607.07914*, 2016.
- [12] Hermon Jonathan. Frogs on trees ? *arXiv preprint arXiv:1609.08738*, 2016.
- [13] H. Kesten and V. Sidoravicius. The spread of a rumor or infection in a moving population. *The annals of probability*, 33(6):2402–2462, 2005.
- [14] Harry Kesten and Vladas Sidoravicius. A shape theorem for the spread of an infection. *Annals of Mathematics*, pages 701–766, 2008.
- [15] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. American Mathematical Soc., 2009.
- [16] Thomas M Liggett, Roberto H Schonmann, and Alan M Stacey. Domination by product measures. *The Annals of Probability*, 25(1):71–95, 1997.

- [17] Mathew D Penrose and Agoston Pisztora. Large deviations for discrete and continuous percolation. *Advances in applied probability*, pages 29–52, 1996.
- [18] S.Y. Popov. Frogs in random environment. *Journal of Statistical Physics*, 102(1):191–201, 2001.
- [19] A. Telcs and N.C. Wormald. Branching and tree indexed random walks on fractals. *Journal of applied probability*, 36(4):999–1011, 1999.